It has produced a wide range of free information sheets and five freephone numbers for recorded information and advice, including legal and financial information (0800 318771/2/3/4/5). There is a wide network of active local branches, some of which provide services for people with dementia and their carers.

Further reading

Twigg J, Atkin K. *Carers perceived.* Buckingham: Open University Press, 1994. (Includes two fascinating accounts, one of the influences of feminism, kinship obligation, the New Right, and the disability lobby on the debate about carers, and the other on the role of general practitioners and consultant physicians.)

Parker G, Lawton D. *Different types of care, different types of carer: evidence from the general household survey.* London: HMSO, 1994.

Green H. *General household survey 1985: informal carers.* London: HMSO, 1988.

Brotchie J, Hills D. *Equal shares in caring.* London: Socialist Health Association, 1991.

1 Goldberg D. *Manual of the general health questionnaire.* Windsor: NFER-Nelson, 1978.
2 Carers National Association. *Who cares? Perceptions of caring and carers.* London: Carers National Association, 1996.
3 Department of Health. *Employers and carers.* London: DoH, 1993.
4 Department of Health and Social Services. *Growing older.* London: HMSO, 1981.
5 House of Commons Social Service Committee. *Community care: carers. Social Service Committee 5th report (1989-90).* London: HMSO, 1990.
6 Warner N. *Better tomorrows. Report of a national study of carers and the community care changes.* London: Carers National Association, 1995.

Statistics Notes

Interaction 1: heterogeneity of effects

Douglas G Altman, John N S Matthews

This is the 24th in a series of occasional notes on medical statistics

In several types of study we may want to examine the consistency of an observed relation across two or more subgroups of the individuals studied. For example, in a clinical trial we might want to know if the observed treatment difference is the same for young and old patients or for different stages of disease at presentation. In an epidemiological study we might want to know whether the observed relation between an exposure and disease is different among smokers and non-smokers

In such cases we are interested in examining whether one effect is modified by the value of another variable. This may be viewed as the examination of the heterogeneity of an observed effect, such as treatment benefit in a clinical trial, across subsets of individuals. The statistical term for heterogeneity of this type is interaction; the medical concept of synergy is the same thing.

While it may well be of interest to look for heterogeneity of effect, this is not always wise. In a controlled trial there are numerous subgroups which might be compared by splitting the patients according to sociodemographic or clinical categories at the start of the trial. In addition, for continuous variables such as age or blood pressure there are many ways of creating groups. Exploratory examination of many such subgroups is almost certain to throw up some spurious significant interactions, and in practice we cannot tell if a specific interaction is real or spurious. For example, in a randomised controlled trial comparing dexamethasone phosphate with placebo for preventing neonatal respiratory distress syndrome, the researchers found unexpectedly that the overall beneficial effect of the active treatment was present only in female infants.[1] Further studies would be needed to confirm the finding (or not). The refutation of such unexpected observations is common, and indeed in this case the finding was not replicated in further studies.[2]

Likewise, we can investigate the interaction between any pair of variables in a regression model. With 10 variables there are 45 such potential interactions and much scope for being misled. So, although we do not necessarily believe that all effects are truly independent, in many cases it is reasonable not to examine any possible interactions. For example, Pocock *et al* found a negative association between tooth lead concentrations and IQ (intelligence quotient) in children aged 6.[3] Exploratory analysis revealed a strong association among boys and little association among girls. They were rightly cautious in their interpretation as there had been no prior hypothesis about such an effect.

By contrast, when there is a specific prior suspicion of the existence of a particular interaction it is perfectly reasonable and desirable to examine it. A common example already mentioned is the interest in a possible difference of risk between smokers and non-smokers. For example, a study of Danish porcelain painters found that the adverse effects of cobalt exposure on lung function were more severe among non-smokers than smokers.[4]

Results of tests for interactions are likely to be convincing only if they were specified at the start of the study. In any study that presents subgroup analyses it is important to specify when and why the subgroups were chosen. Studies which present analyses without such justification can be difficult to interpret. For example, Penttinen found a significant excess of ischaemic heart disease in relation to back pain in farmers aged 30-49 and a non-significant difference in the opposite direction among those aged 50-66.[5] He did not explain why this age division was made, nor did he note that there was no relation when the two age groups were considered together. Studies where subgroup definition has been guided by the data, for example concentrating on males born in October,[6] should be based on statistical tests that account for any multiple comparisons that have been made[7] and should be scientifically sensible; even then they should be treated with scepticism until confirmed in subsequent studies.

Problems of interpretation are exacerbated by incorrect analysis. We consider right and wrong ways to examine possible interactions in two subsequent statistics notes.

ICRF Medical Statistics Group, Centre for Statistics in Medicine, Institute of Health Sciences, PO Box 777, Oxford OX3 7LF
Douglas G Altman, *head*

Department of Medical Statistics, University of Newcastle, Newcastle upon Tyne NE2 4HH
John N S Matthews, *senior lecturer in medical statistics*

Correspondence to:
Mr Altman.

BMJ 1996;313:486

1 Collaborative Group on Antenatal Steroid Therapy. Effect of antenatal dexamethasone administration on the prevention of respiratory distress syndrome. *Am J Obstet Gynecol* 1981;141:276-87.
2 Crowley P, Chalmers I, Keirse MJNC. The effects of corticosteroid administration before preterm delivery: an overview of the evidence from controlled trials. *Br J Obstet Gynaecol* 1990;97:11-25.
3 Pocock SJ, Ashby D, Smith MA. Lead exposure and children's intellectual performance. *Int J Epidemiol* 1987;16:57-67.
4 Raffn E, Mikkelsen S, Altman DG, Christensen JM, Groth S. Health effects due to occupational exposure to cobalt blue dye among plate painters in a porcelain factory in Denmark. *Scand J Work Environ Health* 1988;14: 378-84.
5 Penttinen J. Back pain and risk of fatal ischaemic heart disease: 13 year follow up of Finnish farmers. *BMJ* 1994;309:1267-8.
6 Helgason T, Jonasson MR. Evidence for a food additive as a cause of ketosis-prone diabetes. *Lancet* 1981;ii:716-20.
7 Bland JM, Altman DG. Multiple significance tests: the Bonferroni method. *BMJ* 1995;310:170.